Preprint
UCRL-JC-150494

A Markov-Chain-Monte-Carlo-Based Method for System Identification

R. E. Glaser, C. L. Lee, J. J. Nitao, and W. G. Hanley

This article was submitted to
International Modal Analysis Confernce-XXI, Kissimmee, FL, Feb. 3-6, 2003

October 22, 2002

U.S. Department of Energy

Lawrence
Livermore
National
Laboratory

A Markov-Chain-Monte-Carlo-Based Method for System Identification

Ronald E. Glaser ‡, Christopher L. Lee ¶, John J. Nitao °, and William G. Hanley ‡
‡ Electronics Engineering Technology Division; ¶ New Technologies Engineering Division;
° Geophysical and Environmental Technologies Division
Lawrence Livermore National Laboratory
7000 East Ave.
Livermore, CA 94550

Abstract
This paper describes a novel methodology for the identification of mechanical systems and structures from vibration response measurements. It combines prior information, observational data and predictive finite element models to produce configurations and system parameter values that are most consistent with the available data and model. Bayesian inference and a Metropolis simulation algorithm form the basis for this approach. The resulting process enables the estimation of distributions of both individual parameters and system-wide states. Attractive features of this approach include its ability to: 1) provide quantitative measures of the uncertainty of a generated estimate; 2) function effectively when exposed to degraded conditions including: noisy data, incomplete data sets and model misspecification; 3) allow alternative estimates to be produced and compared, and 4) incrementally update initial estimates and analysis as more data becomes available. A series of test cases based on a simple fixed-free cantilever beam is presented. These results demonstrate that the algorithm is able to identify the system, based on the stiffness matrix, given applied force and resultant nodal displacements. Moreover, it effectively identifies locations on the beam where damage (represented by a change in elastic modulus) was specified.

Introduction
Computational simulations are used to model physical processes in order to predict the behavior or characterize the nature of a given system. To do this they must be consistent with both physical laws and observational data. Acquired data can be integrated into a simulation after inverting the system model (also referred to as the forward model) and identifying the system parameters. The solution to this inversion problem yields the system state which then can be used to predict system behavior. The inversion problem is complicated by issues regarding non-linearity, state-space dimensionality, under/over determined systems, noisy and dependent data, etc. Hence, a strict inversion is rarely possible. In fact, without appealing to unrealistic simplifying assumptions, the severity of these issues often makes classical optimization algorithms ineffective at estimating (i.e., inverting) those system parameters which most consistently correspond to the observed data and the forward model. Moreover, conventional solutions provide little insight into the degree of uncertainty associated with the inversion result.

This work describes an effort to develop and validate a system identification and characterization algorithm that incorporates vibration signature measurements, numerical simulation models of mechanical and structural systems, and a-priori system knowledge to provide an optimal system representation of the actual structure. The solution approach is based upon a general inference paradigm called the Stochastic Engine (SE) [1,2]. This new methodology combines disparate types of observational data and process simulations to produce a consolidated body of knowledge indicating those configurations and system parameter values which are most consistent with the available data and models (i.e., system identification). Bayesian inference and a Metropolis simulation algorithm form the basis for the approach. The resulting procedure enables the estimation of distributions of both individual parameters and system-wide states, and their likelihood of occurrence. Since a posterior distribution on the state space is produced, this approach is capable of yielding quantitative measures of the uncertainty of the generated estimates. This provides the basis for: i) the objective assessment of competing estimates when the available information isn't sufficient to definitively identify the system state, and ii) the propagation of uncertainty in modeling results through to follow-on predictions.

This class of models has the added advantage of being able to follow a system through its lifetime as it changes or ages. Specifically, vibration based signatures of an existing mechanical or structural system can be periodically incorporated into the previous simulation design model to reflect incremental physical change in the actual system. The resulting model can then be compared to the original design model to identify any physical changes in the actual system. For example, these changes may be due to an earthquake which can damage a building, or repeated takeoffs and landings which can fatigue an airframe over time. A critical characteristic of the methodology is its robustness to variety of types of information degradation. Both noise degraded and missing data are considered.

Stochastic Engine Methodology
The Stochastic Engine combines simulations with observations using Bayesian inference that is realized

through a Markov Chain Monte Carlo (MCMC) sampling algorithm. This is accomplished by using the generated samples to estimate a posterior probability distribution indicating those configurations of the system which are most consistent (i.e., most probable) with the existing data and models.

The method begins with the synthesis of a "base representation" of system configurations (i.e., states) and the specification of a prior distribution, _, defined across these states. A search algorithm, employing a form of importance sampling, is used to efficiently traverse the state space. For each visited state, forward simulators are used to predict values of measurable parameters that are then compared to corresponding measurements to determine the likelihood that the given state produced the observed data. These comparisons drive an MCMC algorithm which effectively estimates the posterior distribution over the state space. This information allows one to determine those system configurations that are closest to the true (but unknown) state of nature. Unlike classical inversion techniques that generate a single best-case deterministic solution, this method produces a posterior probability distribution defined over the range of possible solutions - a stochastic inversion of the system of interest. This facilitates post-processing decision analysis and needs-based experimental planning. Moreover, due to its forward processing scheme (i.e., the simulators are all forward models), the methodology is generally applicable to highly non-linear, poorly constrained, multi-dimensional problems.

This methodology places a prior distribution, _, on the set of all possible configurations. Specifically, a Markov chain, P, is constructed with transition probabilities between states such that its limiting distribution is the prior, _. In statistical terms, P samples _. This sampling process is then modified to account for the likelihood of the observed data. The result is a new Markov chain, R, which samples the posterior distribution, __

It should be noted that the Engine includes all available data, model information and inherent error in the system to produce probability distributions identifying likely system configurations or behavior, and quantifying the potential improvement provided by new data. Even when conventional inversion and analysis methods are able to address complex problems, they provide only a single "best" answer, throwing away much of the information and precluding other likely possibilities. On the other hand, the SE allows continuous integration of new data into the analysis, improving understanding and reducing uncertainty.

A Simple Forward Model: Fixed-Free Cantilever Beam

As an illustrative example, we apply the Engine to the simplest of mechanical models: a fixed-free cantilever beam. The SE algorithm is used to identify the configuration of the beam in terms of its stiffness. The uniform, linearly elastic cantilever beam has n equal, homogeneous beam-type elements. We allow each element to have an elastic modulus that is one of three types: (i) type 0 (nominal), (ii) type 1 (abnormal), and (iii) type 2 (abnormal).

Type 0 represents the nominal stiffness of an "unflawed" beam of the given construction. If there exists defects within the beam, it is represented by an element with a type 1 or type 2 elastic modulus. Since each element has one of three possible elastic moduli, and there are a total of n elements, it follows that there are a total of 3^n possible states, or configurations, of the beam, of which exactly one represents an unflawed beam; the remaining 3^n - 1 configurations characterize a beam with at least one flawed element. In the results presented here, we consider the case of n = 10 elements.

We will also restrict the number of possible configurations by assuming that there are at most two flawed elements present in the beam at any one time. As a consequence, the number of possible configurations is reduced from 3^n to $1 + 2n + 4n(n - 1)/2 = 2n^2 + 1$. For n = 10, this is a reduction from 59049 to 201. This restriction is a matter of convenience: to facilitate a short simulation run time and thereby allow a large number of investigative examples.

Static forces applied at selected nodes are specified as the input to the system. The resulting nodal deflections are taken to be the observed output. In the examples, the beam is restricted to two-dimensional, in-plane motion. The static forces are applied normal to the beam in the in-plane direction. Therefore, the displacement at a node has dimensionality d = 2, corresponding to a vertical component and an angular (twisting) component. A sequence of m = 3 independent static forces are applied for a given experiment. The equation of motion for the j^{th} force applied to a beam having configuration i is

$$\mathbf{K}^{(i)}\,\mathbf{X}_j^{0(i)} = \mathbf{F}_j, \tag{1}$$

where $K^{(i)}$ is the nd x nd stiffness matrix determined by the aggregated attributes of the combination of flawed and unflawed elements which comprise the i^{th} configuration, and $\mathbf{X}_j^{0(i)}$ is the nd x 1 theoretical (i.e., mean) displacement vector induced by the nd x 1 applied force vector F_j. The configuration, i, is identified by the number, location(s), and type(s) of flawed elements.

To implement a particular simulation, an actual configuration of the beam is specified, denoted $i = i^*$. Specific force vectors F_j are assumed, and simulated vectors of measured displacements $\mathbf{X}_j^{(i^*)}$ are generated based on the equations of motion and a well defined noise model,

$$\mathbf{X}_j^{0(i^*)} = \left[\mathbf{K}^{(i^*)}\right]^{-1}\mathbf{F}_j, \quad j = 1, \ldots, m, \tag{2}$$

$$X_{jt}^{(i^*)} = X_{jt}^{0(i^*)}\left(1 + \sigma Z_{jt}\right), \quad t = 1, \ldots, nd. \tag{3}$$

For simplicity of notation we define the displacement data,

$$\mathbf{D}_j = \mathbf{X}_j^{(i^*)}. \tag{4}$$

Here Z_j denotes a random vector of independent standard normal variates, and _ > 0 is a parameter. Thus, the measured (i.e., noisy) displacements vary from their theoretical means in a proportional and component-wise

independent and identically distributed fashion. Moreover, π has an easy interpretation: it is the coefficient of variation (i.e., ratio of the standard deviation to the mean) for each displacement component.

Markov Chain Parameterization. We begin by describing our *a priori* knowledge of the sample space of all possible configurations. For our simple model, we will assume there can be no more than 2 flawed elements. There are therefore $N = 2n^2 + 1$ different configurations, of which: (i) 1 consists of no flaws ("level 0"), (ii) $2n$ consist of one flaw ("level 1"), and (iii) $2n(n - 1)$ consist of two flaws ("level 2"). Let us assign total prior probability λ_q to level q:

$$\lambda_0 + \lambda_1 + \lambda_2 = 1, \quad 0 < \lambda_q < 1. \tag{5}$$

For example, by setting each λ_q equal to 1/3 we would be assuming we have no *a priori* knowledge to favor any one of the flaw levels. If we use $\lambda_0 = 0.1$, $\lambda_1 = 0.8$, $\lambda_2 = 0.1$, we would be assuming a particular prior belief that favors the existence of a single flaw. The prior distribution puts probability at each possible configuration. For simplicity we will assume the probability is spread uniformly within each flaw level. Hence,

$$\begin{aligned}
\pi_0 &= \lambda_0 \pi \\
\pi_i &= \tfrac{\lambda_1}{2n} \pi \text{ for each configuration } i \text{ in level 1} \\
\pi_i &= \tfrac{\lambda_2}{2n(n-1)} \pi \text{ for each configuration } i \text{ in level 2.}
\end{aligned} \tag{6}$$

In real applications there may be justification for apportioning prior probability unevenly among configurations within a level, but in our example we will assume uniformity.

We want to construct a Markov chain $P = (p_{ij})$ which samples π. The chain starts at some state and moves from state to state according to transition probabilities p_{ij}, where

$$\sum_{j=1}^{N} p_{ij} = 1, \quad 0 \le p_{ij} < 1, \quad 1 \le i \le N, \quad 1 \le j \le N, \tag{7}$$

and $N = 2n^2 + 1$ is the total number of states. If Y_t denotes the state at time t then we say the chain has limiting distribution π if

$$\lim_{t \to \infty} P\left(Y_t = j \mid Y_0 = i\right) = \pi_j \text{ for any } (j, i). \tag{8}$$

Once a P is constructed to sample π, the MCMC algorithm can be applied to use the likelihood of the observed data to modify P to produce a new chain R which samples π. In our example the data would consist of measured displacements at selected nodes. The likelihood, L_i, of the data assuming a given configuration i is the joint probability density function of the measurements for this configuration.

We will use a special case of the MCMC methodology popularized by Mosegaard [3,4]. Its validity requires the chain P to be reversible, i.e.,

$$\pi_i p_{ij} = \pi_j p_{ji} \text{ for all } i, j. \tag{9}$$

The chain $R = (r_{ij})$ defines the Mosegaard transition probabilities as follows

$$\begin{aligned}
r_{ij} &= p_{ij} \min\left(\tfrac{L_j}{L_i}, 1\right), \quad j \ne i \\
r_{ii} &= 1 - \sum_{j \ne i} r_{ij}
\end{aligned} \tag{10}$$

The chain R makes its transitions based on proposals from the prior sampling chain P. It accepts the proposed state with certainty if the likelihood of the data for the proposed state is greater than that of the current state. If however the likelihood of the proposed state is less, it accepts the proposed state with probability equal to the ratio of the likelihoods, and remains at the current state with probability one minus this ratio. In this way the chain samples the posterior distribution, moving from state to state based upon relative likelihood and prior assumptions, and lingering in the neighborhoods of states having high likelihood.

Our goal, then, is to construct a reversible Markov chain P with limiting distribution equaling the prior π, i.e., a chain which satisfies equation (9). We are motivated by the general strategy to concentrate prior transition probabilities P on neighboring states, so that the posterior chain R tends to explore states within a close range, moving somewhat slowly, almost methodically, and toward modes of the limiting distribution π. In this spirit we use the following general transition rules for P:

- transitions from level 0 to level 1: each state in level 1 is an equally likely destination
- transitions within level 1: can keep location but change type, or retain type but change location to either adjacent position (wrapping around to the opposite end of the beam if necessary)
- transitions from level 1 to level 2: keep level 1 type and location, and add an extra (location, type) at random among the $2(n - 1)$ possibilities
- transitions within level 2: keep flaw locations and change one of the types, or keep one (location, type) at random and choose a new (location, type) at random among the $2(n - 2)$ possibilities
- transitions from level 1 to level 2: keep one (location, type)
- transitions from level 0 to level 2: do not allow
- transitions from level 2 to level 0: do not allow

To be precise, define

$$\begin{aligned}
a_{qr} &= P\left(\textit{transition from level } q \textit{ to level } r\right) \\
\theta_q &= P\left(\textit{type only change within level } q\right)
\end{aligned} \tag{11}$$

and

$p_{00} = a_{00}$, where the state with no flaws has index 0
$p_{0j} = \frac{1 - a_{00}}{2n}$ for each state j in level 1
$p_{i0} = a_{10}$ for each state i in level 1
$p_{ij} = a_{11}\theta_1$ for each type only change within level 1

$p_{ij} = a_{11}\frac{1-\theta_1}{2}$ for each location shift within level 1

$p_{ij} = \frac{a_{12}}{2(n-1)}$ for each allowable transition from level 1 to level 2

$p_{ij} = a_{22}\frac{1}{2\theta_2}$ for each type only change within level 2

$p_{ij} = a_{22}\frac{1-\theta_2}{4(n-2)}$ for each allowable transition within level 2 which introduces a new (location, type)

$p_{ij} = \frac{a_{21}}{2}$ for each allowable transition from level 2 to level 1. (12)

We have imposed the following constraints by construction:

$$\sum_r a_{qr} = 1, \quad q = 0, 1, 2$$
$$a_{02} = a_{20} = 0. \tag{13}$$

The reversibility requirements (9) impose the following additional constraints.

$$1 - \frac{\lambda_1}{\lambda_0} < a_{00} < 1 \; and \; a_{00} \geq 0,$$
$$1 - \frac{\lambda_1}{\lambda_2} < a_{22} < 1 \; and \; a_{22} \geq 0,$$
$$a_{11} = 1 - \frac{\lambda_0}{\lambda_1}(1 - a_{00}) - \frac{\lambda_2}{\lambda_1}(1 - a_{22}) \geq 0. \tag{14}$$

The Likelihood Function. Recall that our example model fixes a particular configuration *i** as the actual configuration that yields the observed displacement data. We assume that the identity of *i** is unknown, although generally restricted to one among the $N = 2n^2 + 1$ possible configurations with up to two flaws, and proceed to use MCMC to sample the posterior distribution over these possibilities. The algorithm continually makes likelihood comparisons between the current and proposed configurations to direct the exploration of the state space.

The likelihood L_i of configuration *i* is a measure of the plausibility of the data under the assumption that the actual configuration is *i*. The normal independent multiplicative error model (3) gives the expression,

$$L_i = \prod_{j=1}^{m}\prod_{t=1}^{nd} \frac{1}{\sigma\left|X_{jt}^{0(i)}\right|\sqrt{2\pi}} \exp\left(-\frac{1}{2\left[\sigma X_{jt}^{0(i)}\right]^2}\left[D_{jt} - X_{jt}^{0(i)}\right]^2\right), \tag{15}$$

where $\mathbf{X}_j^{0(i)} = \left[\mathbf{K}^{(i)}\right]^{-1}\mathbf{F}_j$ denotes the theoretical (mean) displacement vector for a beam with configuration *i* subjected to force vector F_j.

Example Simulations. We have applied the Stochastic Engine to a number of scenarios involving the single cantilever beam, and have been able to explore a variety of topics including:

- sampling of the posterior
- noise in the data
- selection of the prior
- missing data

In subsequent sections, we will describe our efforts to study each of these areas including the models and tests, simulation results and corresponding insights.

Sampling of the Posterior. The objective is to characterize the posterior distribution, _ defined on the collection of possible configurations. Because we have an analytical expression for _, we may compare it to sampled versions of _ obtained by our implementation of the Engine's MCMC methodology and thereby affirm its accuracy. Indeed our simulations have demonstrated that we are properly sampling the posterior. We have typically used simulations of 10 million Markov chain steps, following a burn-in period of 10 thousand steps conducted to "forget" the starting configuration. Without exception we have observed consistent agreement of four significant figures between the simulated relative frequencies and the corresponding theoretical posterior probabilities.

Noise in the Data and Selection of the Prior. The ability of the Engine to ascertain the true identity of the underlying configuration depends primarily on the degree of variation in the observed measurements relative to the theoretical values predicted by the equations of motion. This variation, or noise, is characterized in our modeling approach by the scalar parameter, _, defined in the independent multiplicative error model of equation (3). This parameter is effectively the reciprocal "signal to noise ratio", so that a value of _ = 0.10 implies that the standard deviation of the measured displacement for a given component is 10% of the theoretical (mean) displacement as given by the equations of motion (1). The smaller _ is, the closer the measurements tend to resemble the theoretical values, with complete concurrence assured at the limiting case _ = 0. The parameter _ is propagated in the likelihood function (15) and becomes a major player in defining the posterior distribution _, as well as the Engine's exploration of the posterior by rule (10). In effect, the noise parameter _ and the prior distribution _ compete in characterizing the posterior distribution. If _ is sufficiently small, the observed data will dominate all prior considerations and result in the entire posterior probability being placed upon the actual configuration. On the other hand, by the Bayesian paradigm, the prior beliefs must carry increased weight if the observed data is sufficiently noisy to induce ambiguity. In such noisy scenarios, the resulting posterior is a blend of the prior and the observed data resulting in the suggestion of a number of plausible configurations (i.e., the posterior would assume a multi-modal form with each distinct mode corresponding to a probable configuration).

We illustrate effects of the degree of noise in Figures 1 through 4. Three general noise levels for measured displacements are considered: low noise as represented by _ = 0.01, i.e. a 1% noise to signal ratio; medium noise _ = 0.05; and high noise _ = 0.10. In Figures 1 and 2 the actual simulated configuration has a single flawed element, namely element number 3 of abnormal type 2. In Figures 3 and 4 the actual simulated configuration has two flawed elements, namely a type 2 flawed element at position 3 and a type 1 flawed element at position 4.

These and subsequent figures offer a summary three-dimensional portrayal of the posterior distribution. Recall that there are *N* = 201 possible configurations for a single beam with 10 elements and 2 abnormal flaw types, assuming there are at most two flawed elements in the beam. Each figure shows 56 coordinates, one which represents the zero flaw case, 10 which represent the different flaw locations for the one flaw case, and 45 which represent the different flaw location pairs for the two flaw case. The leftmost corner position represents the zero flaw case. The 10 coordinates along the foreground diagonal to the right of the zero flaw corner position represent the 10 possible single flaw element locations. For example, the coordinate in Figure 1 with the large cone represents a single flaw located at element number 3. The 45 coordinates beyond the foreground diagonal represent the 45 possible flaw location pairs. For example, the coordinate in Figure 3, low noise, with the large cone represents the situation of flawed elements located at position numbers 3 and 4.

The coordinates embody all the possible abnormal flaw types for their respective flaw locations. At each coordinate is a cone whose height equals the total posterior probability, expressed as a percentage, for the configurations associated with the designated flaw location(s). Thus in Figure 1 with low noise we see a cone with height of 100 at the coordinate representing a single flaw at element number 3. This tells us that the posterior distribution has apportioned probability one between the configurations (flaw location, flaw type) = (3,1) and (3, 2). In fact, the posterior has correctly placed probability one at (3, 2) and probability zero at (3, 1) (and everywhere else for that matter). We have decided to sacrifice resolution in depicting uncertainty among abnormal stiffness types in order to have a concise coordinate base. Generally, we have observed that if there is any significant posterior probability at all, it is concentrated at only one of the flaw type possibilities. However, in high noise scenarios, there have been a few cases in which the significant posterior probability at a particular location description has been spread between two different stiffness combinations. In such situations our three-dimensional rendering is admittedly inadequate, and it would be necessary to consult our complete simulation results to obtain sufficient detail to fully characterize the posterior.

The results displayed in Figure 1 demonstrate the overwhelming effectiveness of the Engine algorithm, regardless of the degree of noise, when the prior favors the truth, in this case the fact that the beam contains exactly one flawed element. The prior used here is characterized by _ = (α_0, α_1, α_2) = (0.1, 0.8, 0.1). The posterior has probability essentially unity for the actual configuration, (flaw location, flaw type) = (3, 2), with small amounts of probability placed at two flaw configurations which include (3, 2). For example, the medium noise case (_ = 0.05) posterior distribution puts 0.98 probability at the (actual) one flaw configuration (3, 2), puts probability 0.01 at the two flaw configuration {(3, 2), (10, 1)}, and smears the remaining 0.01 probability among a broad collection of two flaw configurations.

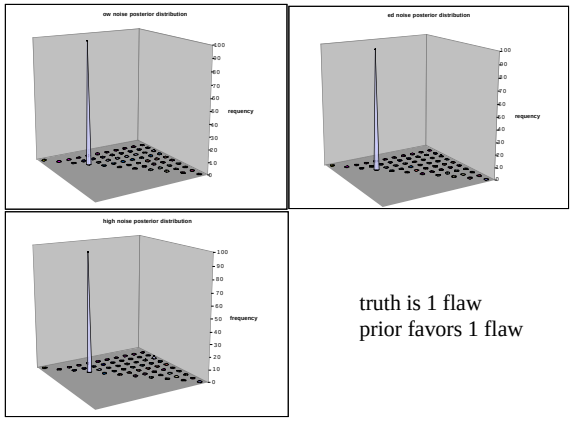

Figure 1. The effect of noise on the posterior distribution for a one flaw scenario in which the prior favors one flaw. The flawed element is at position 3. The noise levels are _ = 0.01, 0.05, and 0.10. Note that in all cases the posterior distribution correctly locates the true state.

In Figure 2, we see the effect of employing a prior which favors configurations contrary to the truth. Once again the one flaw configuration (3, 2) represents truth, but now the assumed prior distribution favors configurations of two flaws: _= (α_0, α_1, α_2) = (0.02, 0.23, 0.75). The observed data in this scenario is not sufficiently compelling to dismiss some two flaw configurations. The Bayesian paradigm conducts a competition between observed data and prior beliefs. The noisier the data, the smaller is its influence in offsetting the strength of the prior. At each noise level in this example, the greatest probability is placed at the actual configuration, (3, 2). In addition, there is significant probability placed at a number of two flaw configurations which include (3, 2). If this were a real-life application, we would conclude that there is clearly a flaw of type 2 at element 3. Moreover, there could be an additional flaw at a location near the end of the beam that might warrant further investigation.

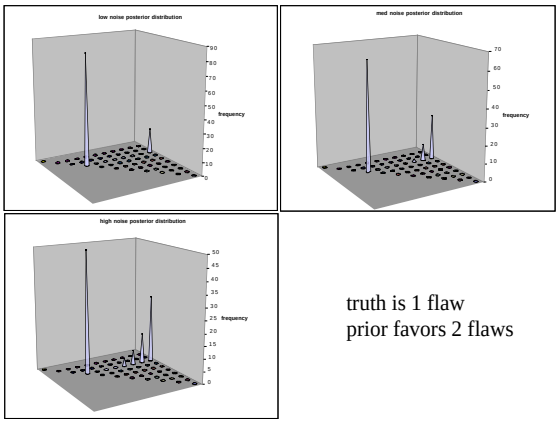

Figure 2. The effect of noise on the posterior distribution for a one flaw scenario in which the prior favors two flaws. The flawed element is at position 3. The noise levels are _ = 0.01, 0.05, and 0.10. Note that the posterior distribution

always indicates that a flaw is present at position 3, but for higher levels of noise the presence of an additional second flaw receives increasing amounts of probability.

Figures 3 and 4 deal with example scenarios in which the actual number of flaws is two, namely the configuration {(3, 2), (4, 1)}. The results are similar to those of the single flaw scenarios shown in Figures 1 and 2. The data overwhelms the prior and the actual configuration is discovered in the low noise setting, while there exists increased posterior uncertainty for increased noise. The effect of a misdirected prior, in this case a prior favoring a single flaw (Figure 4), is to put significant posterior probability at a single flaw subset of the truth, (3, 2) in the high noise case.

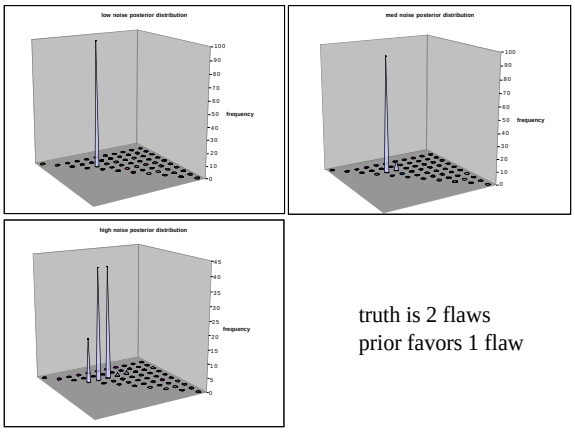

Figure 3. The effect of noise on the posterior distribution for a two flaw scenario in which the prior favors two flaws. The flawed elements are at positions 3 and 4. The noise levels are _ = 0.01, 0.05, and 0.10. Note that posterior becomes increasingly diffuse in a neighborhood of the true configuration as noise levels increase.

These examples essentially demonstrate that the Engine is capable of determining the unknown actual flaw configuration in the low noise situation, and will formulate a posterior distribution in noisy situations that heavily supports the actual configuration but which allows for some alternative plausible configurations (near the true state), motivated to some extent by one's prior beliefs.

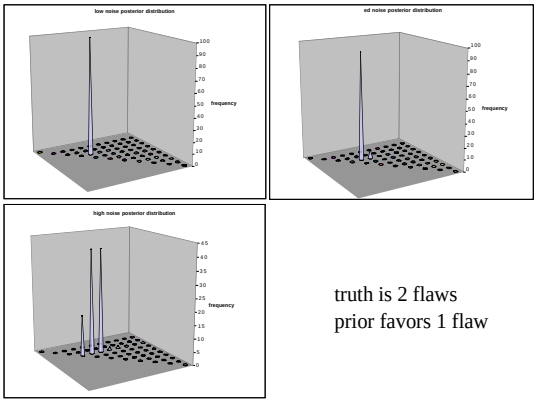

Figure 4. The effect of noise on the posterior distribution for a two flaw scenario in which the prior favors one flaw. The flawed elements are at positions 3 and 4. The noise levels are _ = 0.01, 0.05, and 0.10.

Missing Data. In real life applications, for reasons of economy, convenience, accident, or design, data may not necessarily be collected or available at each node. For the Stochastic Engine, missing data does not in general present an impasse. The posterior distribution can still be sampled and estimated. The only drawback is that the posterior distribution tends to display increased uncertainty as the amount of missing data increases. This is not surprising, as there is less information from which to specify a distribution. In this situation, the calculation of the posterior by MCMC sampling, using (10), merely requires a modification of the likelihood function, (15), such that the product includes only those terms for which measured data is available.

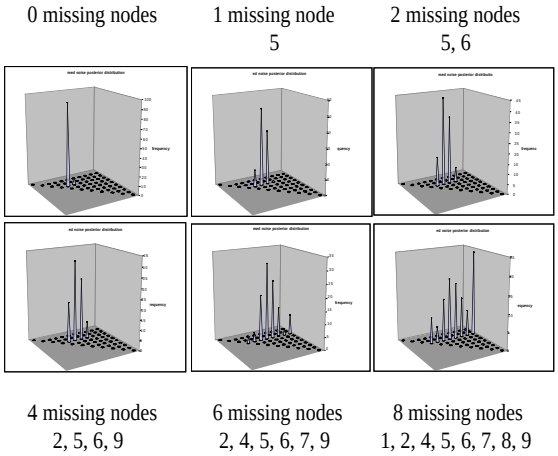

Figure 5. The effect of missing data on the posterior distribution for a low noise (_ = 0.01) two flaw scenario in which the prior favors two flaws. The flawed elements are at positions 3 and 4. The nodes without data are as noted.

We have examined the relationship between the quantity of observed data and the quality of prediction, as measured by the spread of the posterior distribution, for the two flawed element scenarios depicted with full data in Figure 3 of the previous section. We consider incomplete data of the form in which the *dm* = 6 measurements (for the *d* = 2 dimensions and *m* = 3 applied forces) at a particular node are either complete or totally absent. Our results are shown in Figures 5, 6, and 7 for low, medium, and high noise, respectively. In the low noise setting there is negligible loss of precision when 4 or fewer of the 10 available nodes are not measured, and respectable performance with 6 missing nodes. In the medium and high noise settings, however, there is an immediate degradation in the ability to correctly identify both actual flaws: although flaw (location, type) = (3, 2) is consistently and correctly identified, significant posterior probability is placed at configurations away from the other actual flaw, (4, 1). Despite this confusion, for the most degraded test case (i.e., 8 missing nodes and high noise __= 0.10), analysis of the posterior distribution reduces the space of potential location configurations from 56 to 8, an 85.7% reduction in the space of potential hypotheses.

We have conducted alternative simulations with the same amounts of missing data, but with different selections of missing nodes. The results are virtually identical to those shown here. It appears that the number of missing nodes is more important than their actual locations.

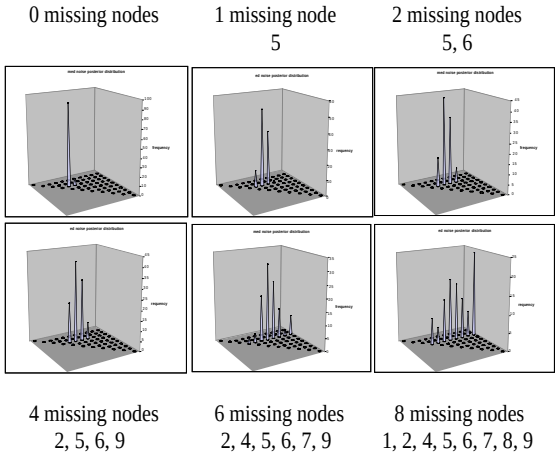

Figure 6. The effect of missing data on the posterior distribution for a medium noise (_ = 0.05) two flaw scenario in which the prior favors two flaws. The flawed elements are at positions 3 and 4. The nodes without data are as noted.

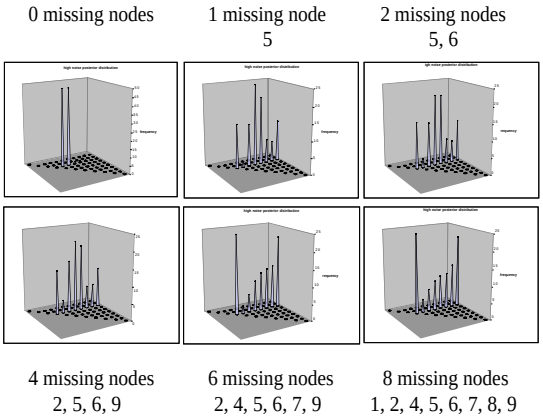

Figure 7. The effect of missing data on the posterior distribution for a high noise (_ = 0.10) two flaw scenario in which the prior favors two flaws. The flawed elements are at positions 3 and 4. The nodes without data are as noted.

Conclusions

A stochastic simulation methodology for the identification of mechanical and structural systems has been presented. This methodology, called the Stochastic Engine, is based upon Bayesian inference and is implemented via a Markov Chain Monte Carlo algorithm. Using a finite element model of a uniform fixed-free cantilever beam as the forward model, the algorithm is shown to identify probable configurations of the beam by determining the stiffness of each beam element through an identification of its elastic modulus. This identification is accomplished through the static deflection of the linearly elastic beam. The system inputs are static forces applied to the nodes of the beam. Each beam element may have one of three elastic moduli values: a nominal or 'unflawed' value, or one of two 'flawed' values. Results show that the methodology successfully calculates posterior probability distributions across the possible configurations of elastic moduli of the beam elements. This distributional information in turn identifies the location of the 'damage,' *i.e.,* identifying the elements with the 'flawed' elastic moduli. Furthermore, the methodology was shown to be robust to the presence of noise in the input and output data and in cases where input and output data sets were incomplete.

In this paper, the output displacements were generated from a 'virtual' experiment (numerical simulation). However it is critical to note that measured response data, as input and/or output, from a real system could be used and incorporated into the finite-element forward model.

Acknowledgments

This work was performed under the auspices of the U.S. Department of Energy by the UC, Lawrence Livermore National Laboratory under contract W-7405-ENG-48.

References

[1] Aines, R.L. et al. *The Stochastic Engine Initiative: Improving Prediction of Behavior in Geologic Environments We Cannot Directly Observe*, UCRL-ID-148221.

[2] Newmark, R.L. et al. *Stochastic Engine: Direct Incorporation of Measurements Into Predictive Simulations.* URCL-JC 145116. Proceedings of the International Groundwater Symposium "Bridging the gap between measurement and modeling in heterogeneous media" Berkeley, CA March 25-28, 2002.

[3] Mosegaard, Klaus (1998) Resolution analysis of general inverse3 problems through inverse Monte Carlo sampling. *Inverse Problems*, v. 14, pp. 405-426.

[4] Mosegaard, Klaus, and Camilla Rygaard-Hjalsted (1999) Probabilistic analysis of implicit-inverse problems. *Inverse Problems,* v. 15, pp 573-583.